Plasticity of the human auditory cortex due to discrimination learning of speech features

H. Menning[1], S. Imaizumi[2], R. Hayashi[2], and C. Pantev[1]

[1]Biomagnetism Center, Institute of Experimental Audiology, University of Münster,D-48129 Münster, Germany; [2]Dept. of Speech and Cognitive Sciences, University of Tokyo, 113- 0033 Tokyo, Japan

1 Introduction

Does learning of non-native speech features lead to a reorganization of language-specific areas in the human brain? It is well known that we perceive and produce speech through the filter of our native language. While infants are able to discriminate most phonetic categories in the first months of life, they already develop prototypical phoneme representations of their native language in the first years of life [1]. Once established, these phoneme categories form a stable base for the consecutive speech perception and speech production [4]. For instance, Finnish and Estonian adults perceive their native phoneme prototypes more sensitively than non-native phonemes [7]. Japanese listeners have difficulties in distinguishing /l/ and /r/ and perceive both phonetic categories as one [2]. However, it has been shown that an intensive frequency discrimi-nation training can enhance the representational areas of the trained frequency [6]. It is assumed as well, that discrimination of non-native speech contrasts can be learned through intensive training [3,5]. Winkler et al. [8] demonstrated that Hungari-ans, who learned in adulthood to speak fluently Finnish, also developed cortical memory representa-tions for Finnish phoneme categories, which do not exist in their native language.

The accurate perception of non-native phonetic categories can be measured by means of an automatic, preattentive change detection component of the Auditory Evoked Response (AER) called the Mismatch Negativity (MMN) and its magnetic counterpart, the Mismatch Field (MMF) [7,8]. The phonemic memory traces revealed by the MMN/MMF seem to be language specific, as there is an enhancement of this component for native but not for non-native phoneme categories. The aim of the present study was to examine if an intensive discrimination training of very small differences of non-native speech contrasts affects the discriminability performance of these features and also affects the brain's automatic change detection response as reflected by the MMF.

2 Methods

In the German language the duration of a phoneme is not a crucial discriminative feature. In other languages the relative duration of a vowel or consonant can be sometimes the decisive feature for differentiating between meaningful words. For instance, in English, the difference between "heed" and "hid", "bid" and "bead" or "wooed" and "wood" is encoded only by the length of the vowel. In languages like Finnish or Hungarian every short vowel has a corresponding long vowel and every long vowel a short counterpart. In a similar way the duration of some consonants is used in a short and long variant. The Japanese language has a restricted amount of possible syllables. This led to a high amount of homonymous syllables in Japanese so that meaningful words are often differentiated only by the length of a vowel or consonant when context is not available. For the present study meaningful Japanese words were generated of natural spoken words at the Department of Speech and Cognitive Sciences of Tokyo, differing in 8 degrees of duration of a vowel or a consonant (Fig. 1). For measurements, the stimuli were lowpass-filtered at 4.5 kHz and the on- and offset of the vowel segment was corrected with a cosine-like slope to eliminate clicks and steep flanks. All stimuli parameters were kept constant except the length of a vowel or a consonant. The first (long) and the last (short) duration always formed one separate perceptual category for Japanese listeners, corresponding to a meaningful word, but not for German listeners, who did not associate any meaning, neither with the long, nor with the short variant. All 8 difference stages of the pairs of stimuli (anni-ani, itte-ite, kiyo-kyo and kiyou-kyou) were presented to 10 native German subjects for discrimination training. Each training session was accompanied by a discrimination test without feedback. Before and after the training period the neuromagnetic responses to 3 different stages of consonant and vowel duration were recorded in a magnetically shielded room with a 37-channel biomagnetometer (Magnes, BTi).

2.1 Discrimination training

Subjects were trained on 10 consecutive days for 1.5 hours according to the stair-case method with visual feedback. Similar to the oddball paradigm in MEG measurements, a frequent standard stimulus pair was interspersed with rare deviant stimulus pairs. The standard stimulus pair consisted always of two equal stimuli with longest duration. The deviant stimulus pair consisted of the stimulus with longest duration and one of 7 stimuli with a shorter duration depending on the discrimination performance reached in the training. The order of occurrence was randomized. The training started at the highest durational difference and continued with progressively smaller durational differences until the smallest difference to the standard stimulus was reached. Each training performance was cross-validated with the performance of a discrimination test without feedback (Fig. 2).

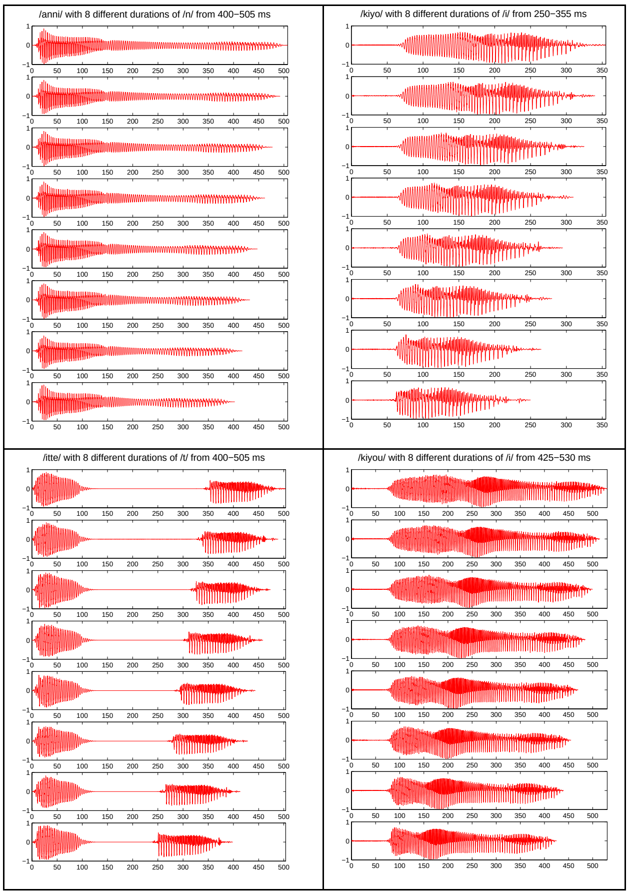

Figure 1**:** *Left: The duration of the consonants in /anni/ and /itte/ is varied in 8 steps of 15 ms. Right: The duration of the vowel /i/ in /kiyo/ and /kiyou/ is varied in 8 steps of 15 ms. In MEG measurements the difference stages 8, 6, and 4 of /anni/ and /kiyo/ were tested. In the discrimination training all 8 differences were trained and tested.*

After each training session subjects were reinforced by informing them about their best performance in that training. A small financial reward was announced for the best performance in all trainings.

2.2 MEG measurements

For MEG recordings of the MMF an oddball paradigm was adopted. Each recording consisted of 2 sessions, one session with 3 different stages of vowel duration (kiyo 8-6-4) and another session with 3 different stages of consonant duration (anni 8-6-4). For each condition 3 runs were grand-averaged. The deviant probability was set to 15% with a randomized ISI of 900 ± 100 ms. The intensity of each stimulus was set to 70 dB SL above each subject's measured detection threshold. Recordings were carried out from the left auditory cortex while stimulated contralaterally.

3 Results

3.1 Discrimination training

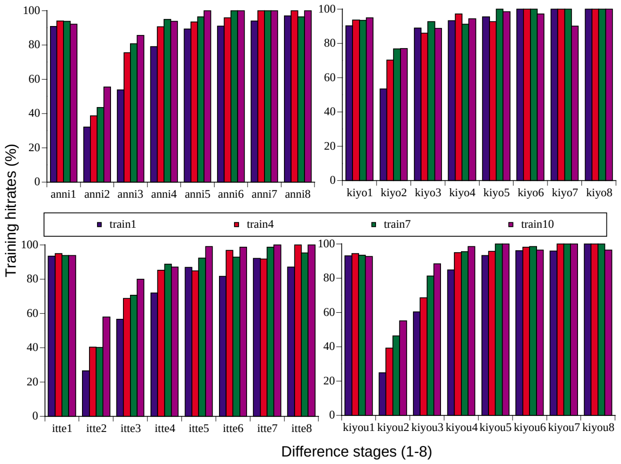

Figure 2: *Hitrates for each of 8 difference stages in every third training with feedback. The discrimination performance increases exponentially in the small difference stages until the end of training. The high differences are differentiated easily at an early time point whereas the performance in small differences increases until the end of training.*

Improvements in the training occurred mainly in the small difference stages 2-4 until the end of training. In large difference stages 5-8 the best performance was reached early and maintained until the end of training (Fig. 2). These findings suggest that there is a categorical boundary at difference stages 4 and 5 between the long and the short version of the stimuli. These results are corroborated by the performance obtained in tests without feedback before and after training sessions.

3.2 MEG measurements

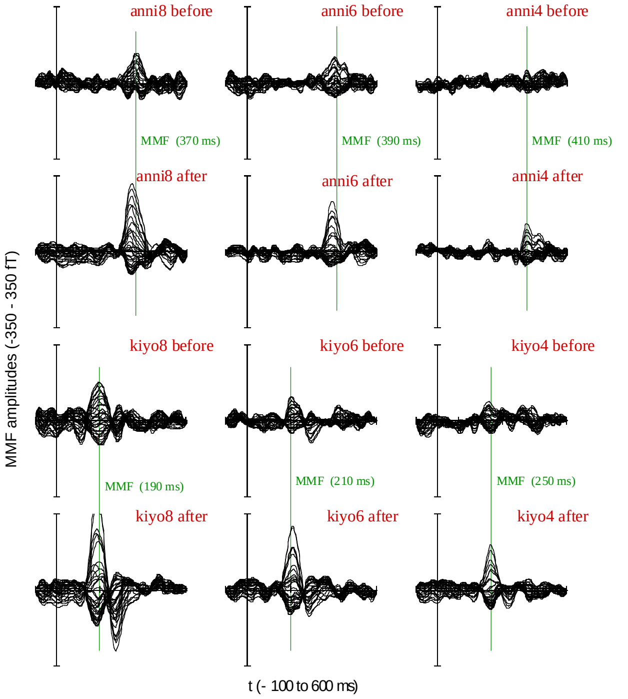

Figure 3: *From left to right: Grand-averaged MMF waveforms from the largest to the smallest difference for /anni/ (top) and /kiyo/ (down) demonstrated on a single subjects data.*

For evaluation of the MMF the grand-averaged responses to the standards were subtracted from grand-averaged responses to deviants. A source analysis based on a single moving dipole model in a sphere was performed. The coordinates estimated here were used to calculate the source strength in a fixed dipole model.

The MMF for /anni/ is recorded from 370 to 450 ms after stimulus onset and from 170 to 250 ms for /kiyo/ (in both cases 150-230 ms after the point of uniqueness of the stimuli). An increase of the amplitudes of the MMF after the training is demonstrated for a single subject in Fig. 3. In grand-averages of all subjects (Fig. 4) a significant increase of the amplitudes of the MMF was found after training for anni8 ($p<0.05$). For anni6 and anni4 the differences were not significant, but for kiyo8 and kiyo6 the differences were highly significant ($p<0.01$), whereas for kiyo4 the difference tended to significance. In the conditions with the large differences (anni8 and kiyo8) a slight decrement of the latency was also observed. The increment in the responses to duration differences in vowels showed a clearer and profounder MMF as to consonant differences (Fig. 4).

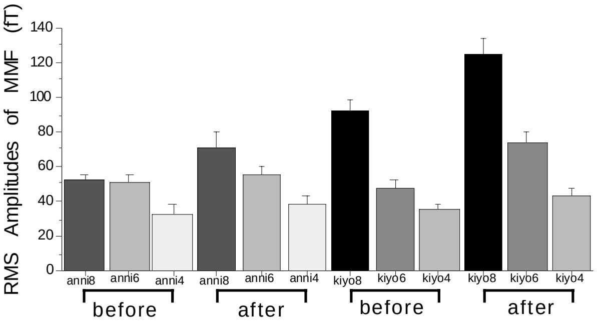

Figure 4: *From left to right: Grand-average of the RMS values (mean, s.e.m.) of all subjects from the largest to the smallest difference (8-6-4) for the consonant condition in /anni/ (left) and the vowel condition in /kiyo/ (right).*

4 Discussion

The obtained results indicate that phonemes differing only in small durational differences are learned to be discriminated after relatively short training time and a performance improvement up to 120% is reached in the smallest difference after the last training compared to the performance before the first training. A similar effect occurs on the neurophysiologic level. The preattentively processed phoneme contrasts, as reflected by the MMF, appear to require more of the neuronal resources after the training. Even small differences are perceived easier and result in a higher activation of the underlying neuronal networks. Speech-related plasticity seems to occur as a consequence of discrimination learning. This implies that the cortical representations of language features can be remodeled dynamically by intensive behavioral training even in adulthood. The effects of amplitude increment and latency decrement is similar to frequency discrimination learning as reported earlier [6]. Intensive training improved not only the discrimination performance of the subjects but changed also the neural representation of speech sounds as reflected by the automatic change detection response, the MMF. Assumed that the different length of phonemes is learned to be perceived as different phonemic categories by the German listeners, the shift of the category boundary to smaller differences achieved in the training can be hypothesized to be due to phonemic processing. Further studies have to investigate if this processing differs substantially from the processing of durational differences in native language or non-speech material with the same parametric characteristics.

Acknowledgements

This work was supported by the "Deutsche Forschungsgemeinschaft" (Pa 392/6-3).

References

1. Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R.: "Development of language specific phoneme representations in the infant brain". *Nature Neuroscience*, **1**, 351-353 (1998).
2. Goto, H.: "Auditory perception by normal Japanese adults of the sounds 'L' and 'R' ". *Neuropsychologia* **9**, 317-323 (1971).
3. Kraus, N., McGee, T.J., Carrell, T.D., King, C., Tremblay, K., & Nicol, T.: "Central auditory system plasticity associated with speech discrimination training". *J Cogn Neurosci,* **7**, 25-32 (1996)
4. Kuhl, P.K.: "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not". *Perception and Psychophysics* **50** (2), 93-107, 1991.
5. Lively, S.E., Logan, J. S., and, Pisoni, D. B.: "Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories". *J Acoust Soc Am* **94** (3), 1242-1255 (1993).
6. Menning, H., Roberts, L.E. & Pantev, C.: "Plastic changes in the auditory cortex induced by intensive frequency discrimination training". *NeuroReport*, **11** (4), 817–822 (2000).
7. Näätänen, R., Lehtokoski, A., Lennes, M. Cheour-Luhtanen, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik., J., Sinkkonen, J., & Alho, K.: "Language-specific phoneme representations revealed by electric and magnetic brain responses". *Nature* **385**, 432-434 (1997).
8. Winkler, I., Kujala, T, Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R.J., Näätänen, R.: "Brain responses reveal the learning of foreign language phonemes". *Psychophysiology*, **36,** 638-642 (1999*).*